

Bibliometria vs *peer review*: è possibile la sperimentazione di nuovi percorsi?

Paola Galimberti, Università degli Studi di Milano

paola.galimberti@unimi.it

Abstract

La brusca accelerazione determinatasi con l'operatività dell'Agenzia nazionale ANVUR, prima, e il varo della valutazione pluriennale della ricerca VQR, poi, ha ri-attivato la discussione e l'attenzione sulle metodiche utilizzate, e utilizzabili, per realizzare un esercizio di valutazione ad ampia scala in grado di rendere comparabili settori, tipologie di prodotti e modalità di accrescimento e di diffusione delle conoscenze scientifiche. A una generale richiesta di disponibilità di informazioni strutturate ha fatto seguito un processo diffuso e partecipato di critica alle diverse proposte metodologiche, processo che ha tuttavia rappresentato un'opportunità di approfondimento e di condivisione circa le possibili soluzioni.

In particolare la discussione si è manifestata all'interno delle aree di scienze sociali e umane (HSS secondo l'acronimo inglese) dove, accanto alla condivisa critica alla bibliometria, è emersa una domanda di bilanciamento rispetto alle eventuali distorsioni della *peer review*, o per facilitare un meccanismo di *informed peer review*. Tra le opzioni emerse, anche nell'ambito del variegato lavoro per l'individuazione dei criteri da parte dei panel tematici/disciplinari, un ruolo particolare è stato rivestito dalla classificazione delle riviste e da quella degli editori, tema che ha visto in campo fautori (molte società scientifiche) e detrattori (in particolare alcuni analisti). I rating e ranking oggi disponibili, che sono il primo prodotto di questa analisi, costituiscono un primo parziale tentativo di qualificare l'analisi in quei settori dove l'assenza o la scarsa robustezza di dati bibliometrici e le probabili distorsioni nel processo di *peer review*, rischiavano di generare un esito scarsamente credibile. Ma forse il dato più interessante che è emerso dall'esercizio è una riflessione sulla modalità con cui queste valutazioni possono essere fatte: chi sono i titolari ad esprimere i giudizi, quali parametri e indicatori utilizzare, come si governano le dinamiche, che uso può essere fatto della documentazione raccolta, in particolare rispetto a possibili diverse destinazioni e alla salvaguardia della condivisione nella comparabilità dei giudizi. Il tema è particolarmente caldo in alcune discipline dove l'impatto non sempre coincide col riconoscimento da parte delle, spesso ridotte, comunità scientifiche o dove esiste un variegato rapporto tra compiti "divulgativi" e compiti "scientifici".

Il paper affronta questa materia rappresentando un'esperienza concreta realizzata (il progetto è però ancora in corso) come attività interateneo di classificazione che ha permesso a un gruppo di atenei partecipanti di creare una base informativa comune, interdisciplinare sulle aree 10-14, partendo però dalle esigenze e dalle specificità delle singole realtà, realizzando un osservatorio che favorisca lo scambio dei dati fra atenei diversi e la comparazione fra le diverse strutture.

Questa attività va colmare un vuoto di strumenti e di informazioni per le aree 10-14 che ha reso fino ad ora gli Atenei autoreferenziali nella valutazione di queste aree e permette in prospettiva una serie di considerazioni e politiche di indirizzo dell'attività scientifica per quanto riguarda le discipline umanistiche¹.

¹ Considereremo umanistiche le aree CUN dalla 10 alla 14, ben consci del fatto che le differenze a livello di SSD sono molto marcate e che alcuni settori hanno un comportamento molto più simile a quello delle scienze dure, fra questi la psicologia, la filosofia della scienza o l'economia.

Valutare nelle aree umanistiche: il complesso rapporto fra qualità e quantità

Il tema della valutazione della ricerca, delle metodiche da impiegare, dei suoi utilizzi e delle sue possibili conseguenze ha popolato negli ultimi tempi quotidiani, blog, convegni e tavole rotonde. Motore di discussioni, critiche e riflessioni è stato l'avvio della Valutazione della Qualità della ricerca (VQR) 2004-2010, il secondo esercizio di valutazione nazionale², che permetterà al Ministero, sulla base dei risultati, di distribuire una quota dell'FFO a partire dal 2013 disponendo di dati più strutturati ed aggiornati rispetto a quelli finora utilizzati per tale scopo. La VQR 2004-2010 ha preso avvio il 7 novembre 2011³ e si concluderà a metà del 2013 e coinvolgerà obbligatoriamente università statali e enti di ricerca vigilati dal MIUR, più altre strutture che hanno chiesto volontariamente di essere sottoposte alla valutazione.

Mentre l'idea della necessità della valutazione dell'università appare largamente condivisa e ormai accettata da tutte le comunità scientifiche, meno condivise sembrano essere le finalità⁴ e ancor meno le modalità attraverso le quali essa deve essere realizzata. L'agenzia Nazionale per la Valutazione della Ricerca (ANVUR) ha dovuto predisporre, in pochissimo tempo, una macchina credibile in grado di associare gli strumenti ai fini dichiarati, dovendo in molti casi svolgere un compito di indirizzo e di giustificazione concettuale delle scelte proposte e, in alcuni casi, di sostituzione rispetto a informazioni, competenze e "tools" non sempre disponibili. In pratica una macchina che, anche in ragione delle scelte fatte sulla quantità e rappresentatività dei prodotti da considerare, permettesse di valutare con metodiche e indicatori differenti a seconda delle aree disciplinari una enorme quantità e varietà di lavori di ricerca, con numeri che avrebbero potuto mettere in crisi sistemi ben più rodati quali quello australiano o britannico⁵. Nonostante gli ostacoli la VQR ha preso avvio, e benché i più scettici sottolineino il suo carattere di ipoteca sulle valutazioni future (anche individuali), questo primo passo ha generato riflessioni e discussioni importanti per una evoluzione delle pratiche valutative nell'ambito della ricerca.

Come avviene in altri sistemi della ricerca anche in Italia la responsabilità della valutazione è stata affidata a panel di area (i Gruppi di esperti della valutazione - GEV) che hanno definito, anche con il contributo delle società scientifiche i criteri secondo i quali i lavori dei ricercatori italiani verranno valutati.

² Il primo esercizio di valutazione nazionale è stato la Valutazione triennale della ricerca (VTR) 2001-2003.

³ Valutazione della qualità della ricerca 2004-2010 Bando di partecipazione (7 novembre 2011)

http://www.anvur.org/sites/anvur-miur/files/bando_vqr_def_07_11.pdf

⁴ *"La valutazione della ricerca non deve essere percepita come un meccanismo burocratico o censorio, ma come preziosa opportunità per mettere a fuoco aspetti nevralgici della performance delle strutture di ricerca..."* Linee guida per la valutazione della ricerca VTR 2001-2003 http://vtr2006.cineca.it/documenti/linee_guida.pdf; *"Tutte le università dovranno ripartire da zero. E quando la valutazione sarà conclusa, avremo la distinzione tra researching university e teaching university. Ad alcune si potrà dire: tu fai solo il corso di laurea triennale. E qualche sede dovrà essere chiusa."* S. Benedetto, intervista a Repubblica 4.02.2012

<http://ricerca.repubblica.it/repubblica/archivio/repubblica/2012/02/04/laurea-doc.html>

⁵ La VQR 2004-2010 riguarderà 195240 lavori di ricerca e 68237 ricercatori, numeri impressionanti rispetto agli altri esercizi nazionali

Dalla lettura dei documenti dei GEV delle varie aree⁶ appare evidente la difficoltà di rendere comparabili aree e settori con differenti tipologie di prodotti, differenti comportamenti rispetto alle scelte delle sedi editoriali, differenti modalità citazionali e di lavoro e variabile composizione delle comunità di riferimento.

Molti dei problemi legati agli strumenti e alle metodiche da applicare per la valutazione della ricerca riguardano in particolare le aree cosiddette HSS (corrispondenti alle aree CUN 10-14) che in Italia hanno cominciato solo negli ultimi anni (alcuni settori solo in epoca molto recente) ad occuparsi di indicatori quantitativi e qualitativi, e a fronte di una scarsa esperienza hanno potuto contare su un'ampia letteratura sul tema (anche in senso critico) e su molte esperienze portate avanti negli altri paesi a livello europeo e internazionale⁷ e quindi evitare di prendere strade già percorse e interrotte da chi le ha precedute.

Per tradizione e anche per una serie di motivi contingenti – cronica mancanza di dati, molteplicità dei canali di comunicazione dei risultati della ricerca difficilmente tracciabili dai database citazionali, differenti modalità e tempi citazionali, uso della lingua nazionale – le scienze umane hanno sempre considerato la “peer review” l'unico strumento di valutazione possibile per certi ambiti disciplinari. Certamente la “peer review” è stata applicata a vari livelli alla valutazione di progetti, persone, gruppi e lavori di ricerca, ma non sempre i criteri utilizzati (dalla scelta dei revisori, alla modalità di revisione, alla ricerca del consenso nel caso di pareri discordi) sono stati resi noti in maniera trasparente. La brusca accelerazione dei processi legati alla valutazione determinatasi con l'operatività dell'ANVUR ha provocato anche l'accelerazione delle procedure di accreditamento delle sedi editoriali (collane, riviste, case editrici) e ha avviato una vivace discussione, non priva di toni polemici, sulle modalità di conduzione della “peer review” presso le sedi editoriali nazionali e sul senso dell'utilizzo e affiancamento di indicatori quantitativi nelle aree 10-14.

Dovendo attuare valutazioni ad ampia scala, che coinvolgono strutture e gruppi di ricerca, la “peer review” non pare uno strumento utilizzabile in forme e modalità tali da garantire la confrontabilità⁸, né il meccanismo di gestione, oneroso in termini di tempi e costi, sembra poter assicurare per queste aree una sostenibilità a lungo termine. Sempre di più quindi si è affermato (questo a dire il vero anche per le aree 01-09) l'idea dell'impiego della *informed peer review*, cioè di una revisione dei pari in cui al valutatore sono resi noti alcuni elementi in grado di supportarne il giudizio o, eventualmente, di correggerlo.

Per le scienze dure la “peer review” viene normalmente affiancata, o preceduta, dagli indicatori bibliometrici che però nelle scienze umane non risultano disponibili per una serie di motivi noti, sia legati alle tipologie di lavori di ricerca (saggi, monografie vs. articoli), sia connessi alle modalità di citazione che ne rendono difficile la raccolta (citazioni a pie' di pagina piuttosto che in fondo all'articolo, citazioni a monografie e saggi piuttosto che ad articoli), o ai tempi di citazione che sono molto più lunghi che per le scienze dure. La struttura stessa della comunicazione scientifica nelle scienze umane ha impedito la creazione di database bibliometrici in grado di fornire indicatori quantitativi e di impatto delle pubblicazioni

⁶ Tutti i documenti elaborati dai GEV sono disponibili all'indirizzo: <http://www.anvur.org/?q=content/composizione-dei-gruppi-di-esperti-della-valutazione>

⁷ Wissenschaftsrat (2010) Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, Knaw, Standard Evaluation Protocol 2009-2015, Royal Netherland Academy for Arts and Sciences (2011) Quality indicators for research in the Humanities, Ben Martin et al., Towards a Bibliometric database for the Social Sciences and the Humanities

⁸ Si pensi alla difficoltà di ricondurre giudizi analitici a giudizi quantitativi.

scientifiche, e i lavori e gli studi orientati in tal senso sono rimasti per il momento a livello puramente teorico⁹.

In assenza di indicatori bibliometrici, ma con la stessa necessità delle scienze dure di poter condurre una “informed peer review”, nell’ambito delle scienze umane alla “peer review” degli articoli si è pensato di affiancare altri tipi di indicatori che riprendono il principio secondo il quale il valore del contenuto corrisponde tendenzialmente al valore del contenitore¹⁰.

Rating di riviste: tre diversi esempi

Già da parecchi anni a livello internazionale, europeo e di singoli paesi circolano elenchi di riviste che classificano i journal in fasce di merito (tre o quattro a seconda delle aree scientifiche) sulla base di criteri predeterminati. Queste liste sono nate per vari scopi. Fra i più noti e diffusi, ERIH (European Reference Index for the Humanities¹¹), ora arrivato alla sua seconda versione, era nato inizialmente con lo scopo di documentare la produzione scientifica europea di qualità lasciando ampio spazio alle pubblicazioni in lingue diverse dall’inglese¹²

Il primo elenco di riviste, molto criticato, è stato pesantemente emendato. Siamo alla fase di revisione quasi completata. I criteri di distribuzione delle classi di merito (che prima utilizzavano le lettere A B e C con l’avvertenza che non si trattava di fasce qualitative) distinguono fra riviste nazionali e internazionali che pubblicano ricerche rilevanti, citate e prestigiose a livello nazionale, europeo o internazionale. La “revised list” pubblicata nel 2011 comprende attualmente solo 10 delle iniziali 15 categorie.

Aeres (l’Agenzia per la valutazione della ricerca francese) ha creato una lista di riviste¹³ per le Scienze sociali e umane per fornire ai valutatori « la liste des revues qui, par discipline ou domaine, tiennent une place importante dans la diffusion des résultats de la recherche à l’échelle nationale et surtout internationale ».

L’Agenzia francese in una prima fase (2008) aveva anche suddiviso i titoli in fasce di merito, nelle liste aggiornate invece c’è un’unica fascia di merito che comprende tutte le riviste accreditate.

⁹ Si veda lo studio commissionato da società scientifiche europee, Ben Martin “Towards a Bibliometric Database for the Social Sciences and Humanities –A European Scoping Project”, (2010): In un recente convegno a Venezia lo stesso Ben Martin ha dichiarato che non è ancora ben chiaro se i costi per la costituzione di un database bibliometrico per le scienze umane e sociali sarebbero giustificati dai benefici.

¹⁰ Questo assunto, valido per chi utilizza indicatori riferiti alle riviste (IF SJR SNIP ecc.), risulta smentito dalla distribuzione delle citazioni all’interno dei diversi journal che assai spesso è asimmetrica (pochi articoli che ricevono il maggior numero delle citazioni, molti articoli che non ricevono alcuna citazione) secondo il principio di Pareto. Tuttavia Bonaccorsi (2012) sostiene che l’assegnazione di riviste a classi di qualità fornisce una informazione a priori circa il valore atteso della qualità dei singoli articoli. Nelle riviste con alto indice di impatto è più difficile pubblicare, c’è un maggiore rejection rate e la selezione è più severa, ci si aspetta dunque una qualità più elevata degli articoli pubblicati.

¹¹ <http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities.html>

¹² It is also a unique project because, in the context of a world dominated by publication in English, it highlights the vast range of world-class research published by humanities researchers in the European languages.
<http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities/erih-foreword.html>

¹³ <http://www.aeres-evaluation.fr/Publications/Methodologie-de-l-evaluation/Listes-de-revues-SHS-sciences-humaines-et-sociales>

L'Australian Research Council (ARC)¹⁴ ha deciso di avviare nel 2008 un censimento delle riviste significative per tutte le aree scientifiche elaborando una prima lista di 19500 periodici divisi in 4 classi (A*, A, B, C) poi sottoposta al giudizio delle comunità scientifiche che hanno fatto integrazioni e cancellazioni. Alcuni articoli dimostrano come nel passaggio dalla lista iniziale a quella delle società scientifiche ci siano stati riaggiustamenti anche molto significativi¹⁵.

Alla fine del lavoro di censimento l'ARC ha deciso che la lista non sarebbe stata utilizzata per ERA 2010 (l'esercizio di valutazione australiano)¹⁶. Il ministro per la ricerca e l'innovazione ha accolto le lamentele di molti ricercatori che giudicavano la lista inadeguata e fortemente sbilanciata, sottolineavano i possibili effetti sulle scelte editoriali future che avrebbero soffocato le riviste molto specialistiche e non nel mainstream e soprattutto lamentavano il possibile utilizzo (scorretto) di tale lista per la valutazione dei singoli: "There is clear and consistent evidence that the rankings were being deployed inappropriately within some quarters of the sector, in ways that could produce harmful outcomes, and based on a poor understanding of the actual role of the rankings,"¹⁷.

Le liste di riviste in Italia

L'attività di censimento e classificazione in fasce di merito delle riviste ha preso avvio anche in Italia con una orchestrazione meno centralizzata e organizzata che negli esempi sopra citati partendo, come avviene in molti casi, dal basso. Questo ha fatto sì che a fronte di un lavoro serio e accurato da parte delle singole società scientifiche, i risultati finali siano difficilmente comparabili in quanto i criteri utilizzati, la suddivisione in fasce, l'ampiezza di ciascuna fascia sono state decise in maniera indipendente e difficilmente riconducibile ad una unità.

Da parecchi anni le società scientifiche stanno lavorando sulle liste di riviste, definendo i criteri di inclusione e attribuzione alle classi di merito. Alcune società scientifiche, quelle più orientate verso l'internazionalizzazione, includono nella lista tutte le riviste straniere significative per la disciplina, attribuendo di default alle riviste presenti nei database citazionali (WOS principalmente) la classe di merito più alta (si pensi alla recente lista della Società italiana degli economisti¹⁸).

Altre società scientifiche hanno invece classificato e suddiviso in classi di merito solo le riviste italiane partendo dal presupposto che comunque le riviste di lingua inglese, se di valore, sono già incluse nei database citazionali e sono quindi già dotate di indicatori bibliometrici.

Il risultato di queste attività non sono dei ranking in cui le riviste vengono messe in ordine di punteggio, ma appunto delle classi di merito di ampiezza variabile a seconda dei criteri definiti a priori dalle società scientifiche stesse e in molti casi validati dalle comunità scientifiche di riferimento.

¹⁴ <http://www.arc.gov.au/>

¹⁵ Si veda ad es. l'articolo di J. Lamp, At the sharp end: journal ranking and the dreams of academics, Online information review 33(4) 2009, pp 827-830

¹⁶ http://www.arc.gov.au/era/era_2012/review_of_era10_ranked_outlet_lists.htm

¹⁷ Il ministro aggiunge nella sua comunicazione: Instead of rankings the Australian system will incorporate "journal quality profiles." ... "the removal of the ranks and the provision of the publication profile will ensure they will be used descriptively rather than prescriptively." <http://chronicle.com/article/Journal-Ranking-System-Gets/127737/>

¹⁸ <http://www.siecon.org/online/wp-content/uploads/2012/03/Nota-CdP-13.03.2012.pdf>

Quando si creano strumenti di questo tipo è necessario conoscerne i limiti, non pensare che la qualità del contenitore sia sempre e comunque trasferibile al contenuto, né che queste liste possano essere strumenti oggettivi, che permettono di superare la soggettività e i limiti dei giudizi dei pari¹⁹. Tutti gli strumenti di misurazione quantitativa sono imperfetti, sono comodi perché danno una rappresentazione sintetica e semplificata della realtà, ma non permettono di cogliere i dettagli. A differenza di quanto avviene nei database citazionali, fare una lista di riviste significa valutarne “il profilo scientifico-editoriale secondo parametri dotati di un certo valore e in base a scale di pesi prestabilite. Ne emerge un valore numerico che però serve solo ad esprimere sinteticamente quel lavoro analitico qualitativo e anche con un certo margine di soggettività. Questa non la chiamerei bibliometria, ma analisi diretta qualitativa su parametri dati. Naturalmente, in una simile procedura di valutazione, sono i parametri a dover essere oggetto di discussione preliminare e di consenso, la loro definizione in termini di importanza assoluta e di importanza relativa²⁰”. Ovvero, secondo le parole di Abbattista, i criteri per stilare una lista di riviste non devono prescindere dalla consapevolezza dello scopo a cui tale lista è destinata e della soggettività nella sua compilazione.

Quali parametri nell'analisi bibliometrica delle aree umanistiche e sociali?

Preso atto della difficoltà a coniugare quantità e qualità, vanno anche ricordati gli esercizi che sono stati prodotti in tale ambito, a partire dal documento che più di tutti ha rappresentato, per le società scientifiche e le consulte, una guida e un riferimento nell'elaborazione delle liste di riviste almeno per quei settori delle aree 10-14 in cui la bibliometria non trova applicazione²¹. Tale documento, che va sotto il nome di Bozza Graziosi (6 marzo 2009), fornisce raccomandazioni a società scientifiche e consulte in relazione ai criteri da utilizzare nella formulazione delle proprie proposte di classificazione di riviste ed editori, in particolare i criteri per le riviste. Per le riviste italiane vengono indicate le classi di merito (da A a D) e il peso di ciascuna classe, i criteri di qualità a cui ciascuna rivista valutata deve rispondere. Le raccomandazioni suggeriscono anche possibili modelli di pesi da attribuire alle pubblicazioni presenti nelle riviste nelle diverse classi di merito, nell'ottica di premiare le pubblicazioni nelle fasce superiori e rendere impossibile che l'accumulo di pubblicazioni della fascia più bassa possa portare a valori paragonabili a quelli di chi ha articoli (magari in numero minore) solo nella fascia più alta. Per le riviste straniere si prevede invece una suddivisione in tre classi di merito considerando già meritorio il fatto di pubblicare all'estero.

Molte società scientifiche (soprattutto nelle aree 10 e 11) hanno operato in questi anni seguendo o adattando le raccomandazioni del documento; tuttavia la brusca accelerazione determinata dall'avvio della

¹⁹ Distorsioni legate a giudizi soggettivi e opportunistici sono state pesantemente sottolineate nelle liste di discussione ad esempio in merito alla classifica delle riviste di biblioteconomia e bibliografia, che premia solo uno dei due settori, e trascura il fatto che qualsiasi scienza applicata deve misurarsi costantemente con la professione, che è insieme applicazione dei risultati della ricerca e ricerca applicata.

²⁰ Si veda il recente intervento di G. Abbattista, Riviste umanistiche: ranking e “rankling” Roars, 5 marzo 2012 <http://www.roars.it/online/?p=4872>

²¹ Come vedremo più avanti ogni area ha le sue specificità. All'interno delle singole aree abbiamo settori che ragionano in termini di indicatori bibliometrici come le scienze dure (si pensi alla linguistica o alla filosofia della scienza o alla psicologia e all'economia) e aree che invece sono scienze umane “pure”.

VQR 2004-2010, ha affidato ai GEV in accordo con le società scientifiche il compito di proporre liste “ristrette” di riviste da utilizzare nella procedura di *informed peer review* per le aree umanistiche, facendo perdere di vista la questione dei criteri e attirando critiche e perplessità. Ciascun GEV ha lavorato su classi di merito differenti, operando aggiustamenti e cancellazioni non immediatamente comprensibili o trasparenti. “Tenuto conto del poco tempo a disposizione e che gli elenchi non sono vincolanti per i referee, ci sarà tempo di verificare la robustezza di questo lavoro di rating, di analizzare quanto il valore del contenitore influisca sul valore del contenuto e il tipo di correlazione. Fare degli errori in questa fase, per alcuni settori più nuova che per altri, è normale. Quello che appare invece più preoccupante sono le previsioni di lavoro che fissano i criteri per il futuro e sui quali si dovrebbe riflettere con maggiore attenzione per cercare di evitare errori a cui sarebbe difficile porre rimedio”²².

A conclusione del lavoro di rating dei GEV di area umanistica e nell’ottica di tracciare un percorso che potesse tranquillizzare gli scettici, i critici, i perplessi e le schiere di detrattori, è stato pubblicato a cura del vicecoordinatore della VQR, Andrea Bonaccorsi, il documento “Potenzialità e limiti dell’analisi bibliometrica nelle aree umanistiche e sociali”.

Il documento di Andrea Bonaccorsi vuole essere un documento metodologico di indirizzo per l’attività futura di rating. Esso riconosce tutte le difficoltà di elaborazione di indicatori quantitativi per le scienze umane, mette in rilievo come essi abbiano senso per grandi numeri, archi temporali sufficientemente ampi e con le opportune normalizzazioni e sottolinea il fatto che un primo passo, almeno in via sperimentale, potrebbe essere la creazione di un database bibliografico/bibliometrico con la collaborazione di editori ed istituzioni. Sempre secondo Bonaccorsi le strade da perseguire sono comunque diverse e parallele:

- La candidatura delle migliori riviste italiane(fascia A?) alla indicizzazione in WOS e Scopus con un possibile sostegno di ANVUR²³.
- Il rating delle riviste a partire dalle liste dei GEV, in collaborazione con le società scientifiche e con il contributo ampio e diffuso delle comunità scientifiche di riferimento. Questa seconda strada deve fare tesoro delle esperienze internazionali, tenendo presente che è possibile ricondurre i giudizi esperti a classi di merito. Il documento suggerisce i passi per ottenere liste robuste , primo fra tutti la definizione delle diverse classi di merito, del numero delle classi e delle percentuali di riviste in ciascuna classe, gli intervalli temporali per la revisione delle liste. Particolarmente importante sembra essere la scelta dei valutatori, la composizione dei panel e la formazione del consenso all’interno di ogni panel. Il documento riporta inoltre altre raccomandazioni quali la consultazione online delle comunità scientifiche e la validazione delle liste da parte di revisori esterni.

Benché molto articolato il documento non presenta lo stesso livello di robustezza in tutte le sue parti. Particolarmente debole appare la parte dedicata ai criteri da adottare per le future revisioni dei rating, che riprende in parte le indicazioni della bozza Graziosi che andrebbero senza dubbio meglio definite e articolate anche e soprattutto con l’ausilio di specialisti di scienza dell’informazione ed informatici.

²² Per una breve analisi dei punti critici dei documenti di lavoro dei GEV si veda P. Galimberti Basi dati e scienze umane: quali scelte compiere? Roars 9 marzo 2012 <http://www.roars.it/online/?p=5288>

²³ L’operazione darebbe alle riviste italiane un bollino di qualità, ma non risolverebbe il problema che hanno le Humanities in tutti i paesi, e cioè il fatto che molte citazioni sono a materiali non source.

La valutazione della ricerca in ambito umanistico a livello di atenei

A differenza della Agenzia nazionale che, per valutare le strutture, terrà in considerazione solo l'eccellenza della produzione dei ricercatori italiani²⁴, gli atenei si trovano nella situazione di dover valutare e dare un peso a tutti i lavori prodotti dai ricercatori afferenti ai diversi dipartimenti, e dunque a doversi dotare di strumenti che permettano la comparazione fra le strutture sull'intera produzione scientifica.

Attualmente non esiste in Italia una base di dati che tracci l'intera produzione scientifica italiana nelle aree 10-14²⁵. I dati sono distribuiti in fonti diverse (sito ministeriale, banche dati specialistiche, anagrafi locali, motori di ricerca tipo Google Scholar), ma in molti casi risultano incomparabili perché sporchi, raccolti su archi temporali differenti, con una metadattazione incompleta. Di fatto pensare di effettuare una comparazione fra atenei diversi risulta estremamente laborioso (se non impossibile) in termini di tempi e costi e, data l'incertezza sulla completezza e correttezza dei dati, non dà alcuna garanzia sulla credibilità del risultato. L'unico strumento sembra essere, al momento, il mutuo e consapevole (e volontario) scambio di informazioni, possibilmente minimizzando o interpretando le diversità che derivano da procedure e esigenze tarate per singolo ateneo.

Per agevolare lo scambio dei dati e la comparazione è infatti essenziale che le anagrafi della ricerca locali utilizzino le medesime definizioni per le tipologie dei prodotti della ricerca censiti, (a partire dalla definizione di cosa sia una pubblicazione scientifica²⁶), i medesimi metadati per la loro descrizione, le medesime procedure per la valutazione dei lavori e i medesimi criteri di riferimento.

Partendo dalla comune esigenza di valutare anche in termini comparativi tutta la ricerca svolta nelle discipline umanistiche prive di indici di bibliometrici, un gruppo di quattro Atenei (Milano, Padova, Torino, Bologna) si è dunque accordato per cercare di colmare il vuoto di strumenti e di informazioni per le aree 10-14, ben consci del fatto che tale vuoto tende a rendere gli Atenei autoreferenziali nella valutazione di queste aree. L'iniziativa era stata inizialmente motivata dalla necessità di condividere criteri su quell'area grigia e disomogenea costituita dalla monografie, ma poi il progetto è rapidamente stato indirizzato alla costruzione di una piattaforma comune per lo scambio dei dati e degli strumenti condivisi per la valutazione dei lavori di ricerca (in prima battuta degli articoli) per permettere in prospettiva l'individuazione di percorsi e di considerazioni e politiche di indirizzo dell'attività scientifica per quanto riguarda le discipline umanistiche.

I quattro atenei (a cui di recente si è aggiunto anche quello fiorentino) hanno composizione variabile rispetto alle aree e agli SSD, ma hanno in comune la grande dimensione, la natura "generalista" con ciò garantendo una buona copertura di tutti i settori e costituendo, nel loro insieme, un campione rilevante rispetto alla situazione nazionale.

²⁴ In media 3 pubblicazioni nell'arco temporale 2004-2010.

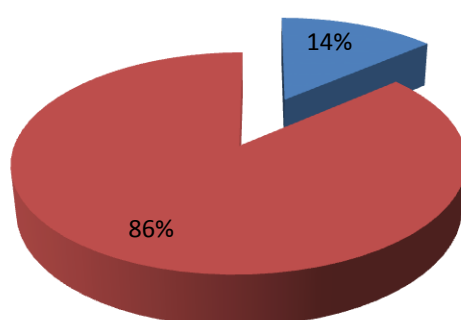
²⁵ Per le scienze cosiddette "dure" ci sono WOS e Scopus con copertura variabile a seconda delle aree.

²⁶ Definizione che a tutt'oggi manca. Abbiamo il parere del CUN sui criteri identificanti il carattere scientifico delle pubblicazioni http://consulta.arte.unipi.it/documenti/Valutazione%20CUN%2024_001.pdf (parere del CUN) 24/02/2010, abbiamo il DM 19 marzo 2010 che definisce le tipologie di pubblicazioni ammesse alle valutazioni e dunque le pubblicazioni ritenute scientifiche http://civr.miur.it/vqr_decreto.html e il DM 17 del 15 luglio 2011 che presenta qualche leggera modifica rispetto al precedente decreto del CIVR http://www.anvur.org/sites/anvur-miur/files/vqr_d.m._n._17_del_15_07_2011_firmato.pdf

Infatti in alcune aree l'insieme dei docenti dei quattro atenei, per numerosità, arriva a coprire il 12% dell'insieme nazionale, in altre invece, come la 14, raggiunge addirittura il 24 % del totale nazionale.

Area	mi	to	pd	bo	totale atenei gdl	totali nazionali
10	188	232	157	122	699	5330
11	136	221	250	264	871	4884
12	207	155	129	203	514	4889
13	63	161	128	295	647	4820
14	71	89	63	300	523	1745
totale atenei gdl	665	858	727	1184	3434	21668

Aree 10-14 gdl rispetto ai totali nazionali



Il progetto prende in considerazione la produttività totale dei ricercatori nelle diverse aggregazioni sotto esame, pesando però questa produttività sulla base del prestigio o reputazione della sede editoriale. Costituisce quindi un buon esempio di differenziazione rispetto all'esercizio VQR (non misura solo l'eccellenza) e dà concreto seguito al come pesare la sede di pubblicazione rispetto ai contenuti e al loro impatto atteso.

I quattro atenei si sono accordati in prima istanza sull'oggetto del processo di valutazione rappresentato da "aggregazioni" di ricercatori (gruppi di ricerca, Dipartimenti, Aree scientifiche) e non da singoli ricercatori. Le aggregazioni saranno definite e concordate di volta in volta a seconda delle finalità della valutazione e saranno identificate "minime comuni aggregazioni" sulla base delle diverse necessità operative. Oggetto primario della valutazione non è l'eccellenza (o qualità) dei singoli prodotti della ricerca, ma l'operosità (o produttività) delle aggregazioni sopra definite (tema rilevante anche in aree scientifiche molto basate sulle contribuzioni individuali), essenzialmente sulla base della reputazione della sede di pubblicazione. Conseguentemente verrà valutata la rilevanza della sede di pubblicazione, da rivedere periodicamente, senza entrare nel merito della qualità intrinseca dei prodotti della ricerca.

Nel condurre la valutazione si è deciso di tenere conto dei differenti stili editoriali nei diversi ambiti del sapere, così come dei differenti gradi e della diversa rilevanza dell'internazionalizzazione della produzione scientifica.

Nella scelta dei criteri per la valutazione sono stati presi in considerazione i parametri, le attività e i risultati delle diverse "agenzie di valutazione" nazionali (es. ANVUR, CUN, CEPR - Comitato di Esperti per la Politica della Ricerca) e, per quanto possibile, internazionali (es. European Research Council, European Science Foundation, ERA, AERES), con attenzione anche per i documenti e rapporti elaborati anche ai fini delle nuove procedure di concorso (es. CUN).

La prima fase del lavoro ha visto, dopo un censimento dei diversi modelli disponibili, la discussione e la stesura di linee guida metodologiche per la classificazione delle riviste, successivamente sono stati identificati i criteri per la classificazione a partire dalla verifica delle condizioni necessarie allo status di rivista scientifica.

L'arco temporale preso in considerazione per i quattro atenei è il 2008-2010. Per questi anni e per tutti i docenti e ricercatori degli atenei sono state estratte dalle rispettive anagrafi le riviste su cui si è pubblicato, per ogni rivista sono stati indicati gli SSD degli autori e la numerosità di contributi per SSD ed è stato creato un elenco comune che riporta gli identificativi internazionali (ISSN, EISSN,) quelli nazionali (ance), Il titolo, i settori e gli atenei che hanno pubblicato su ogni rivista.

A1	ID_NORM								
	A	B	C	D	E	F	G	H	
1	ID_NORM	TITOLO_UFFICIALE_NO RM	ISSN	CD_SSD	N_PRODOTTI_SSD	AREA_CUN	UNI	ISSN/ ISBN	Pe
2	E055017	.ECO L'EDUCAZIONE SO	1972-9995	SPS/09		1 14 - Scienze politiche	UniTo	X	
3	E190942	900. PER UNA STORIA D	2036-8836	SPS/06		1 14 - Scienze politiche	UniPd	X	
4	E190942	900. PER UNA STORIA D	2036-8836	SPS/13		1 14 - Scienze politiche	UniBo	X	
5	E188844	AAA. TAC: ACOUSTICAL	1824-6176	SPS/07		1 14 - Scienze politiche	UniPd	X	
6	E193197	ACHAB. RIVISTA STUDEI	1971-7954	SPS/08		1 14 - Scienze politiche	UniMi	X	
7	E185813	ACOMA. RIVISTA INTERN	1122-6218	SPS/04		2 14 - Scienze politiche	UniPd	X	
8	E185813	ACOMA. RIVISTA INTERN	1122-6218	SPS/05		1 14 - Scienze politiche	UniBo	X	
9	E184808	ACTA HISTRIAE	1318-0185	SPS/03		1 14 - Scienze politiche	UniPd	X	
10	E002014	ACTA PHILOSOPHICA	1121-2179	SPS/08		3 14 - Scienze politiche	UniBo	X	
11	E192154	ACTA SOCIOLOGICA	0186-6028	SPS/09		1 14 - Scienze politiche	UniTo	X	
12	E002435	ACTUEL MARX	0994-4524	SPS/01		1 14 - Scienze politiche	UniPd	X	
13	E201796	ADMINISTRACIÓN & CIDA	1887-5262	SPS/09		1 14 - Scienze politiche	UniMi	X	
14	E193337	AESTHESIS	1751-9853	SPS/08		1 14 - Scienze politiche	UniPd	X	
15	E186125	AETHIOPIA	1430-1938	SPS/13		1 14 - Scienze politiche	UniBo	X	
16	E003869	AFFARI SOCIALI INTERN	0390-1181	SPS/08		2 14 - Scienze politiche	UniTo	X	
17	E003951	AFRICA (ROMA)	0001-9747	SPS/13		12 14 - Scienze politiche	UniBo	X	
18	E003951	AFRICA (ROMA)	0001-9747	SPS/13		1 14 - Scienze politiche	UniPd	X	
19	E185749	AFRICA E MEDITERRANE	1121-8495	SPS/07		2 14 - Scienze politiche	UniPd	X	
20	E185749	AFRICA E MEDITERRANE	1121-8495	SPS/08		1 14 - Scienze politiche	UniBo	X	
21	E185749	AFRICA E MEDITERRANE	1121-8495	SPS/08		1 14 - Scienze politiche	UniTo	X	
22	E199862	AFRICAN HISTORICAL RE	1753-2523	SPS/13		1 14 - Scienze politiche	UniBo	X	
23	E186568	AFRICANA	1592-9639	SPS/13		1 14 - Scienze politiche	UniBo	X	
24	E186568	AFRICANA	1592-9639	SPS/13		1 14 - Scienze politiche	UniTo	X	
25	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/04		1 14 - Scienze politiche	UniBo	X	
26	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/04		1 14 - Scienze politiche	UniMi	X	
27	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/07		1 14 - Scienze politiche	UniBo	X	
28	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/13		17 14 - Scienze politiche	UniBo	X	
29	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/13		1 14 - Scienze politiche	UniMi	X	
30	E183870	AFRICHE E ORIENTI. RIV	1592-6753	SPS/14		2 14 - Scienze politiche	UniMi	X	

Ogni ateneo ha costituito dei panel interni (con composizione simile) a cui è stato affidato, in una prima fase e sulla base dell'elenco comune, il compito di individuare le riviste scientifiche, escludendo la coda di riviste divulgative²⁷. Quando i quattro atenei avranno terminato questa prima fase gli elenchi verranno confrontati e verrà stilato un nuovo elenco che comprenderà solo le riviste individuate come scientifiche. In questa fase le valutazioni dovrebbero coincidere, nel caso però di valutazioni difformi sia all'interno di ogni ateneo che fra i quattro atenei è stata prevista una modalità per arrivare ad un giudizio uniforme.

Come condizioni necessarie allo status di rivista scientifica si sono individuati:

- la presenza di un codice ISSN,

²⁷ L'analisi della "coda" farà parte di altri tipi di riflessione

- la periodicità regolare (almeno annuale),
 - la presenza di un comitato editoriale composto da studiosi del settore di pertinenza delle riviste con riconosciuta qualifica a livello nazionale e che operano presso Università, Enti di ricerca, Soprintendenze e AFAM,
 - la pubblicazione di articoli che riportano ricerche originali,
- Le riviste che non soddisfano tutte e quattro le condizioni elencate entrano a far parte di una “coda” di riviste non scientifiche²⁸.

I criteri per la suddivisione in classi prevedono la creazione di 4 classi di merito A-D, con l’insieme delle riviste di fascia A compreso fra il 10% e il 33%, l’insieme delle riviste di fascia C e D che costituisce almeno il 33% delle riviste e l’insieme delle riviste di fascia D che costituisce il 10% delle riviste classificate.

I criteri per determinare le fasce di merito terranno in considerazione:

- L’utilizzo della peer review,
- La presenza in cataloghi, repertori, banche dati disciplinari, biblioteche di rilevanza nazionale e internazionale
- La presenza di autori stranieri e/o di contributi anche in lingue di verse da quella del Paese nel quale è pubblicata la rivista
- La predisposizione all’indicizzazione: titolo, indici, abstract, parole chiave ecc. anche plurilingue
- La reputazione nella comunità scientifica

Laddove presenti verranno tenuti in considerazione anche i seguenti criteri:

- La presenza o valutazione nelle liste internazionali: ERIH, AERES, ERA
- La presenza in database citazionali
- La disponibilità di IF

Nella prima fase del lavoro, quella in corso, si sta procedendo alla definizione del set di riviste scientifiche.

Una prima analisi dei dati può dare già alcune indicazioni interessanti.

Innanzitutto sono stati estrapolati i titoli senza ISSN che dunque automaticamente vengono esclusi dalla lista comune. Si tratta di circa un 10% dei titoli (tranne che per l’area 13 dove i titoli senza ISSN sono il 5%).

AREA	Totale (titoli/pubb)	1 ateneo (titoli/pubb)	2 atenei (titoli/pubb)	3 atenei (titoli/pubb)	4 atenei (titoli/pubb)
10	1452	1139(2018)	219(969)	80(662)	22 (338)
11	2217	1676(1913)	355(888)	137(672)	49(221)
12	737	496(556)	114(300)	72(241)	56(399)
13	1336	938(2036)	206(1119)	101(1086)	46(321)
14	788	618(951)	111(122)	45(64)	14(24)

²⁸ Non si tratta di un giudizio qualitativo, ma della pertinenza rispetto all’ambito scientifico

In generale si evidenzia un numero molto elevato di titoli presenti in un unico ateneo, titoli a carattere locale o regionale, tuttavia in alcune aree (ancora una volta l'area 13 ad esempio) la condivisione dei titoli (e la concentrazione degli articoli nei titoli condivisi) pare più marcata. Se si associa la compresenza su tutti gli atenei a una sorta di individuazione del core dell'area, risulta evidente che la dimensione delle rispettive aree è ancora troppo grande per trarne indicazioni significative. Oppure, simmetricamente, che vi è un'elevata "dispersione" corrispondente alle specificità e alla pluralità di opportunità di sedi per la pubblicazione che sembrano confermare il trade-off del mondo biologico nel rapporto tra efficienza/produttività e biodiversità. In questo senso la garanzia del pluralismo va coniugata con la presenza di requisiti di qualità/scientificità delle sedi editoriali soprattutto per quanto riguarda le scelte e gli orientamenti futuri²⁹. Ben consci della molteplicità dei target e delle funzioni della ricerca in queste aree (scientifici ma anche divulgativi e culturali in generale) i quattro atenei ritengono che non possa essere attribuito a tutte le sedi editoriali un valore uguale e che il lavoro di censimento e osservatorio comune (con il relativo reciproco controllo) possa portare ad una classificazione verosimile che serva da guida per le future scelte.

Un secondo elemento d'analisi deriva dall'esercizio della *peer review* nei suoi diversi aspetti. Mancano ad esempio indicazioni circa la scelte, sia predefinite ex-ante, sia realizzate ex-post a seguito di esiti non positivi, nella tracciabilità dei processi che portano gli autori all'identificazione delle riviste verso cui indirizzare le candidature. Il tutto, naturalmente, anche in ragione delle diverse opportunità relative alle dimensione e alle caratteristiche dei settori, nonché al grado di "vicinanza" nei confronti dei rispettivi board. Certamente alcune riviste sono palesemente target per i docenti di un unico Ateneo (riviste di Facoltà e/o di Dipartimento), concentrandosi lì la maggior parte dei contributi, sarà compito dei vari atenei riflettere su quali caratteristiche dovranno assumere in futuro queste pubblicazioni cercando di orientarle verso una maggiore apertura ai contributi esterni³⁰.

Un ulteriore elemento d'analisi/riflessione è rappresentato dal confronto del "core" di riviste comuni ai 4 atenei, in alcuni casi molto piccolo come nell'area 14, con i rating fatti dai GEV. Nell'area 12 ad esempio, sulle 56 riviste che costituiscono il patrimonio comune dei 4 atenei 28 sono comprese fra quelle in fascia A, 10 in fascia B, 1 in fascia C e 17 non sono prese in considerazione nel rating.

Conclusioni

Il progetto interateneo è nato da una esigenza di trovare una base comune (che permetta la confrontabilità) per la valutazione della produzione scientifica delle aree 10-14. Gli atenei coinvolti hanno dovuto fare i conti in primo luogo con una assenza di dati funzionali allo scopo. Il primo sforzo è stato quello di definire un set comune di dati a partire dalla definizione degli oggetti da valutare, ponendo quindi le condizioni per una collaborazione multilaterale e, contemporaneamente, dettando delle possibili basi per un lavoro a una scala più generale.

Il progetto del gruppo interateneo è entrato da poco nella sua fase operativa. E' stato preceduto da una lunga fase di gestazione, nel corso della quale le priorità sono state rimodulate. Nato per trovare un

²⁹ Anche per questo motivo si è deciso di prendere in considerazione un arco temporale breve e recente, a partire dal 2008, quando già la questione dei criteri di qualità aveva cominciato a circolare.

³⁰ Una delle FAQ dell'ANVUR per la VQR 2004-2010 e relativa alla presentazione dei prodotti, fissa le caratteristiche che devono avere i WP per poter essere presi in considerazione. Un lavoro analogo potrebbe essere fatto per le riviste di Facoltà o Dipartimento.

sistema comune di pesi e misure per la valutazione delle monografie (e saggi), si è scontrato con una serie di difficoltà nella classificazione delle sedi editoriali³¹ che ha riorientato il lavoro verso una suddivisione delle riviste in classi di merito (rating). Una prima analisi sui dati di partenza comuni ha fatto emergere una forte specificità di sede rispetto alle diverse aree, per cui un core condiviso di riviste (estremamente ridotto) non sembra essere lo strumento più opportuno da utilizzarsi, in particolare se collegato all'analisi dell'intera produzione dei diversi Dipartimenti.

Le motivazioni di natura economica che possono legare un gruppo di ricercatori ad una determinata rivista sono difficilmente determinabili ex post, e l'assenza di una base di dati bibliografico/bibliometrica ostacola una analisi selettiva del rapporto che lega i componenti del board di una rivista ai diversi autori.

Le riviste non sono, in questo momento, l'output principale dei ricercatori delle aree 10-14 (forse con esclusione di parte dell'area 13); tuttavia il lavoro di classificazione sarà utile affinché in futuro, pur nel rispetto della pluralità, nelle scelte delle sedi editoriali si tenga conto di criteri qualitativi condivisi.

Fino ad ora non si è pensato di considerare, pur nei noti limiti della parzialità del fenomeno nelle aree umanistiche, il tema della collaborazione fra gli atenei partecipanti, tuttavia questo resta un ambito da approfondire in futuro, soprattutto per analizzare le reti di relazioni e la loro consistenza.

Il lavoro che il gruppo sta svolgendo e svolgerà mira a colmare un vuoto contingente di informazioni e strumenti. Questa però non deve essere l'unica strada da percorrere. Allo scopo di rendere il processo più democratico e trasparente è necessario orientarsi verso la logica dell'apertura: predisponendo database interoperabili che contengano anche i full-text dei lavori censiti, rendendo trasparenti gli algoritmi bibliometrici e utilizzando forme di peer review aperta.

BIBLIOGRAFIA

ANVUR, Consiglio direttivo (8 marzo 2012), La bibliometria della VQR, http://www.anvur.org/sites/anvur-miur/files/la_bibliometria_della_vqr.pdf

Bonaccorsi A., (2012) Potenzialità e limiti dell'analisi bibliometrica nelle aree umanistiche e sociali. Verso un programma di lavoro http://www.anvur.org/sites/anvur-miur/files/valutazione_aree_umanistiche_e_sociali.pdf

CIRC <http://epuc.cchs.csic.es/circ/que.html>

CNRS http://www.cnrs.fr/comitenational/sections/doc/categorisation37_0911.pdf

Ellison G. (2011), Is peer review in decline?. *Economic Inquiry*, 49: 635–657. doi: 10.1111/j.1465-7295.2010.00261.x

Era 2012 http://www.arc.gov.au/era/era_2012/era_journal_list.htm

KNAW (2010), Standard evaluation Protocol 2009-2015, http://www.know.nl/Content/Internet_KNAW/publicaties/pdf/20091052.pdf

³¹ In primo luogo la mancanza di trasparenza nei processi di revisione dei lavori sottoposti, e le forti differenze presenti non solo presso uno stesso editore ma anche presso una stessa collana

KNAW (2011) Quality indicators for research in the Humanities,
http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20111024.pdf

Lamp J. (2009), At the sharp end: journal ranking and the dreams of academics, Online information review 33(4), pp 827-830

Linee guida per la valutazione della ricerca VTR 2001-2003
http://vtr2006.cineca.it/documenti/linee_guida.pdf

Martin B. et al. (2010) Towards a bibliometric database for the Social Sciences and Humanities – A European Scoping Project. A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF, 8 March 2010,
http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/esf_report_final_100309.pdf

REF 2014 <http://www.hefce.ac.uk/research/ref/panels/members.htm>

ROARS, Return On Academic research, www.roars.it

Taylor, J. (2011), The Assessment of Research Quality in UK Universities: Peer Review or Metrics?. British Journal of Management, 22: 202–217. doi: 10.1111/j.1467-8551.2010.00722.x

Van Gestel R. e Vranken J., Assessing Legal Research: Sense and Nonsense of Peer Review versus Bibliometrics and the Need for a European Approach, 12 German Law Journal 901-929 (2011)

Wissenschaftsrat (2010) Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, <http://www.wissenschaftsrat.de/download/archiv/10039-10>.